

## DOCUMENT RESUME

ED 189 166

TM 800 339

AUTHOR Guay, Poland E.  
TITLE Spatial Ability Measurement: A Critique and an Alternative.  
PUB DATE Apr 80  
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).  
EDRS PRICE MF01/PC01 Plus Postage.  
DESCRIPTORS Adults; \*Aptitude Tests; Cognitive Ability; \*Cognitive Processes; Cognitive Tests; Spatial Perception; Test Construction; \*Test Interpretation; Test Reliability; \*Test Validity; Visual Measures  
IDENTIFIERS \*Purdue Spatial Visualization Test; Spatial Skills; \*Spatial Tests

## ABSTRACT

The construct of spatial ability is discussed and it is suggested that some widely used and cited tests that are called spatial ability tests may not be valid measures of that ability. Instead, their items may be solved using mental processes that are clearly analytical and not spatial in nature. Four studies involving the analysis of subjects' self-reports of processing are cited to show that different individuals employ different processes and that an individual may change processing strategies within the same test. The Minnesota Paper Form Board is given as an example of a test that is likely to be solved by perceptual-analytic and not by spatial processing. The Purdue Spatial Visualization-Rotation Test (PSVR) is described as an example of a test that was designed to minimize analytic processing and processing variability. Reliability and validity data for the PSVR are reported for several studies including analysis of self-reports of item processing, correlations with other tests, sex-difference studies, and supervisor's ratings of job performance. (Author/CTM)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

ED189166

U S DEPARTMENT OF HEALTH,  
EDUCATION & WELFARE  
NATIONAL INSTITUTE OF  
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-  
DUCED EXACTLY AS RECEIVED FROM  
THE PERSON OR ORGANIZATION ORIGIN-  
ATING IT. POINTS OF VIEW OR OPINIONS  
STATED DO NOT NECESSARILY REPRESENT  
OFFICIAL NATIONAL INSTITUTE OF  
EDUCATION POSITION OR POLICY

SPATIAL ABILITY MEASUREMENT:  
A CRITIQUE AND AN ALTERNATIVE

A Paper  
Presented at the 1980 Annual Meeting  
of the  
American Educational Research Association  
April, Boston

by  
Roland B. Guay  
Associate Professor  
Department of Industrial Education  
Purdue University  
and  
Test and Measurement Consultant  
Applied Management Services, Inc.

"PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

R. Guay

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)."

TM 800339

SPATIAL ABILITY MEASUREMENT:  
A CRITIQUE AND AN ALTERNATIVE

Abstract

The construct validity of some widely used and cited spatial ability tests in educational research (e.g., the DAT Space Relations) is suspect since their tasks can be, and often are, solved using mental processes that are clearly analytical and not spatial in nature. Spatial ability measurement is further hampered because the same spatial tasks are often solved by the use of many different analytic and spatial processing strategies. The literature suggests that this construct validity problem needs more serious consideration by researchers. Introduced is a new spatial ability test, the Purdue Spatial Visualization-Rotations Test, that is designed to minimize the analytic processing and processing variability that often confound spatial ability measurement. Limited reliability, validity, and other statistical data are presented on the new test.

SPATIAL ABILITY MEASUREMENT:  
A CRITIQUE AND AN ALTERNATIVE

Roland B. Guay  
Purdue University and  
Applied Management Services, Inc.

Critique

Spatial visualizing ability has been recognized as a distinct and important dimension of human abilities for over half a century. Although spatial ability has never attained the level of recognition afforded verbal and numeric abilities, it has long been recognized as important for success in many occupations and school subjects, such as engineering, drafting, and art. As a result, educational research in these subject areas often includes the spatial variable.

Researchers interested in measuring the spatial variable have dozens of spatial tests to choose from. However, when these so-called spatial tests are carefully analyzed for information-processing strategies, many of them are found to be poor measures of spatial ability because their tasks can be, and often are, solved using non-spatial mental processes. The problem is that several of these poor measures of spatial ability (e.g., the Revised Minnesota Paper Form Board Test and the DAT Space Relations Test) are widely used and cited by educational researchers as measures of spatial ability.

### Spatial Ability

In order to understand the problem, it is helpful to be aware of what is recognized in the field as the critical mental component of spatial ability. In defining spatial ability, Kelley (1928) wrote about the existence of two elements: (1) the perception and retention of visual images; (2) the mental manipulation of these images. Numerous psychologists (e.g., French, 1951; Michael, Guilford, Fruchter & Zimmerman, 1957) since Kelley have also described the components of spatial thinking. Despite the numerous definitions that have been advanced, there is general agreement in the literature today (Bock & Kolakowski, 1973; Harris, 1976; Smith, 1964; Vandenberg, 1975) that it is the mental manipulation involved, and not the perception or retention, which enables a task to measure spatial ability.

In spite of the consensus that spatial ability is more than perception, it is not uncommon for any test using visual stimuli to be cited at one time or other as a measure of spatial ability. The fact that a test utilizes tasks requiring the perception of spatial configurations does not guarantee that it measures spatial ability. An obvious example of this is the Raven's Progressive Matrices. Although at first glance it appears to be spatial in nature, the test loads primarily on the verbal factor (Burke & Bingham, 1969).

Between- and Within-Individual Variation in Processing Strategies

The problem in spatial ability measurement goes well beyond the mislabeling of tests by users. More important is that many tests, which both the authors and publishers claim to be measures of spatial ability, use tasks that can be solved in several different ways. Since the cognitive processing strategy used to solve the task determines what ability the task measures, many so-called spatial tests may be measuring distinctly different abilities for different people. Evidence that many spatial tests may be measuring different abilities is presented by Borich and Bauman (1972) and Price and Eliot (1975). Both these studies found that the correlations between spatial tests are often quite low.

Evidence that many spatial tests can be solved different ways is provided by four studies involving the analysis of subjects' self-reports of processing. Barratt (1953); French (1965); Guay, McDaniel, and Angelo (1978); and Talsma and Wheatley (1980) each found that different individuals often employ different processes in solving the same spatial tasks. Guay et al. (1978) also found that an individual often changes processing strategies within the same spatial test. In terms of measurement, the use of alternate processing strategies would not be as serious a problem if all the different mental processes used were spatial in nature.

However, the self-report studies have found that many of the strategies used are non-spatial.

It is interesting to note that some studies (e.g., Allen, 1974; McGlone & Kertesz, 1973) have found evidence to suggest that males and females prefer different styles of processing spatial tasks, with males favoring spatial strategies and females non-spatial strategies. This difference in preferred processing has been advanced as a potential explanation for the sizable and reliable sex difference in spatial ability favoring males that has been widely substantiated in the literature (e.g., Harris, 1976; Maccoby & Jacklin, 1974).

#### Non-Spatial Processing

The susceptibility of spatial test items to non-spatial processing strategies serves to confound spatial ability measurement. This problem is not new. Smith described the problem in his book published in 1964. Smith observed through his review of spatial tests that many tasks used in tests of spatial ability could be solved using non-spatial processing that is primarily analytical in nature. This analytic processing was described as involving explicit trial-and-error checking of relationships between different parts of a figure. Analytic processing, therefore, is distinctly different from spatial processing, which involves the transformation of visual images. Smith (1964) reported that spatial tests vary in the degree to which they elicit

the use of analytic processing. This notion was supported by the findings reported by Guay et al. (1978) and Talsma and Wheatley (1980).

Example of Poor Spatial Ability Measurement

The analytic processing and processing variability that often confound spatial ability measurement can be illustrated with the Revised Minnesota Paper Form Board Test (MPFB). The MPFB is clearly presented in the test manual (Likert & Quasha, 1970) and the publisher's catalog as a valid measure of spatial visualization ability. A sample item from the test is presented in Figure 1. The examinee

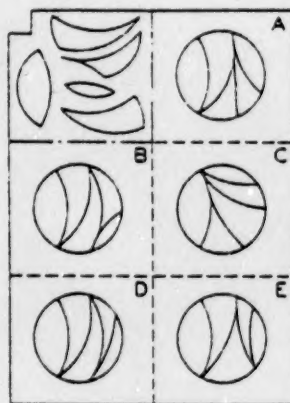


Figure 1. Item from the MPFB Test.

is required to select from the five alternatives the one that shows how the parts given in the top left-hand corner fit together.

Despite what the test authors and publisher say, there is sufficient evidence to suggest that the MPFB is not as valid a measure of spatial ability as claimed. For example, in the Guay et al. (1978) study, which included an analysis of self-reports of information-processing strategies used by 16 university students to solve MPFB Test items, it was found that nearly all the students used non-spatial, analytic processing strategies; and many of these students changed their processing strategies as the MPFB items increased in complexity. For simple items with few parts, most students reported that they found the correct solution by simply comparing the number, shape, and size of the individual parts to those parts making up each figure. For more complex items with more parts, however, many students reported that they developed a strategy of concentrating on one "characteristic part." For the item presented in Figure 1, for instance, these students reported that they concentrated on the smallest part among the five parts given in the top left-hand corner. By doing so, they were able to eliminate alternatives A and D because they obviously did not include this characteristic part. Through close visual inspection, eventually alternatives C and E were also eliminated because their smallest parts, although having the same shape as the characteristic part, were somewhat larger than the characteristic part. The same or similar perceptual-analytic

approaches were reported by Guay et al. (1978) to be used frequently and successfully in taking the MPFB Test. The study by Talsma and Wheatley (1980) found that the MPFB elicited the use of perceptual-analytic processing by all of the 24 seventh- and eighth-grade students participating in the study.

The problem with the posited construct validity of the MPFB can also be observed empirically. For example, in a study of 101 university students by Guay and McDaniel (1978), correlations between scores on the MPFB and scores on three relatively pure-factor spatial tests were found to be quite low (i.e., around .30). In addition, in the same study the expected large sex differences in spatial test performance were found for the three pure-factor spatial tests ( $p < .001$ ), but not for the MPFB Test. Female and male group scores on the MPFB were reported to be almost identical. These data again suggest that the MPFB may be measuring something other than spatial ability for the majority of the people taking the test.

#### Summary

The construct validity problem is not unique to the MPFB Test. Other examples of so-called spatial tests that are poor measures of spatial ability are much too easy to find. In each case, the problem is basically the same, i.e., the susceptibility of the tasks to the use of more than one

processing strategy to solve the tasks, and more importantly, the fact that many of the processing strategies used do not require the mental manipulation of visual images which is considered the critical component of spatial visualization ability.

It is logical that the best measures of spatial ability are those tests which maximize the use of spatial processing strategies while concurrently minimizing both between-individual and within-individual variation in processing. The vulnerability of spatial tasks to analytic processing and processing variability make the interpretation of many spatial test scores difficult and, thus, the use of these scores in research questionable. Based on past and present practice in educational research, there appears to be a need for more serious consideration by researchers for the construct validity of the spatial tests they use and cite.

#### An Alternative

Among the dozens of spatial ability tests available to educational researchers is a new instrument titled the Purdue Spatial Visualization-Rotations Test (PSVR). The PSVR is part of the Purdue Spatial Visualization Test Battery (Guay, 1977) and is included in the ETS Test Collection. The PSVR is not introduced here as the solution to the problem of spatial ability measurement; rather, it is offered as an example of a spatial test that has been designed and observed (Guay

et al., 1978) to minimize the analytic processing and processing variability that often confound spatial ability measurement.

### Description of the PSVR Test

The PSVR is designed to measure a specific type of spatial visualizing ability that requires imagining the movement of three-dimensional figures according to explicit directions. The test has 30 items, a time limit of 20 minutes, and is appropriate for use by individuals of age 13 or older. A sample item is presented in Figure 2. The directions are self-explanatory. The correct answer for the sample item is given on the bottom of the last page.

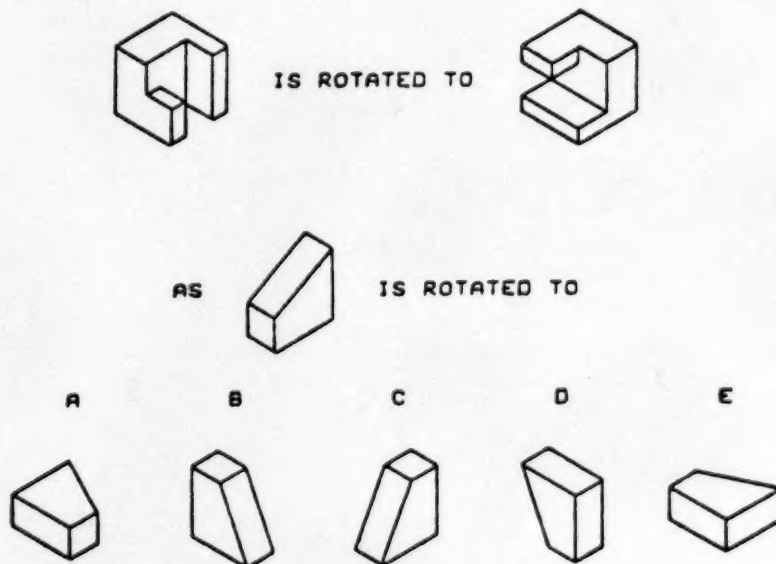


Figure 2. Item from the PSVR Test.

The item shown in Figure 2 is relatively difficult compared to the other PSVR items. Item difficulty is primarily a function of the number of axes on which the depicted object is rotated and the amount of rotation on each axis. For the sample item, the object is rotated 90 degrees on two axes. The PSVR Test starts off with items requiring a rotation of 90 degrees on one axis. These items are followed by items requiring a rotation of 180 degrees on one axis, then items requiring a rotation of 90 degrees on two axes, and, finally, items requiring a rotation of 90 degrees on one axis and 180 degrees on another axis.

The same stimulus object is used to specify the required rotation in all 30 PSVR items (see the top line of the item in Figure 2). The other stimulus object included in the item is different for each item. These other stimuli are either truncated hexahedrons, right circular cylinders, right rectangular prisms, or right triangular prisms.

The administration and scoring of the PSVR Test is not difficult. Both individual and group administrations are appropriate. The test booklet is designed to be reused and has complete instructions with example items. Separate answer sheets used with the PSVR may be either hand or machine scored. In scoring, correction for guessing formulae are not employed. The PSVR score is simply the number of items answered correctly.

### Data on the PSVR Test

Limited data on the PSVR is available from three studies: Guay (1978a), conducted on 217 university students; Guay (1978b), conducted on 51 skilled machine maintenance employees; and Guay and McDaniel (1978), conducted on 101 university students. In regards to the reliability of the PSVR Test, from these studies internal consistency coefficients (KR-20) of .87, .89, and .92 were observed. These coefficients suggest that the PSVR scores may be sufficiently accurate to be used for both group and individual decisions.

The PSVR was developed specifically to serve as a research instrument designed to measure spatial visualization ability. Some evidence for the construct validity of the test is provided in the study by Guay et al. (1978). Through the analysis of self-reports of processing in this study, it was found that the PSVR tasks require spatial processing and not perceptual-analytic processing in order to arrive at the correct solution. Other evidence for the construct validity is provided by the Guay and McDaniel (1978) study. In this study, a correlation coefficient of .61 was observed between the PSVR scores and scores on the Shepard-Metzler Rotations Test (Shepard & Metzler, 1971). Still other evidence for the construct validity of the PSVR Test is provided by data on between-sex differences. In both the Guay (1978a) and Guay and McDaniel (1978) studies, males scored significantly

higher ( $p < .001$ ) on the PSVR than females. According to the literature, this result is expected for measures of spatial ability.

The PSVR also appears to have some potential for predicting performance in various occupations. In the study by Guay (1978b), a concurrent validity coefficient of .36 was observed between scores on the PSVR and supervisor ratings of job performance for the skilled maintenance employees.

Other findings with PSVR Test data are also quite revealing. A significant birth order effect ( $p < .05$ ) and a significant experience effect ( $p < .01$ ) are reported in the study by Guay (1978a). A probe of these results revealed that university students who were born neither first nor last in their family scored significantly higher than students born first or last, and university students who often participated in activities requiring spatial thinking scored significantly higher than students who seldom participated in these same activities.

The Guay (1978b) study reported significant positive correlations ( $p < .001$ ) between the PSVR Test scores and scores on both a mechanical reasoning test and a quantitative reasoning test. The coefficients were .63 and .50, respectively. The tests were the SRA Test of Mechanical Concepts (Stanard & Bode, 1976) and the SRA Verbal Test (Thurstone & Thurstone, 1973).

## References

- Allen, M. Sex differences in spatial problem-solving styles. Perceptual and Motor Skills, 1974, 39, 843-846.
- Barratt, E. An analysis of verbal reports of solving spatial problems as an aid in defining spatial factors. The Journal of Psychology, 1953, 36, 17-25.
- Bock, R., & Kolakowski, D. Further evidence of sex-linked major-gene influence on human spatial visualizing ability. American Journal of Human Genetics, 1973, 25, 1-14.
- Borich, G., & Bauman, P. Convergent and discriminant validation of the French and Guilford-Zimmerman spatial orientation and spatial visualization factors. Educational and Psychological Measurement, 1972, 32, 1029-1033.
- Burke, H., & Bingham, W. Raven's Progressive Matrices: More on construct validity. Journal of Psychology, 1969, 72, 247-251.
- French, J. The description of aptitude and achievement tests in terms of related factors. Psychometric Monograph No. 5, 1951.
- French, J. The relationship of problem-solving styles to the factor composition of tests. Educational and Psychological Measurement, 1965, 25(1), 9-28.

Guay, R. Purdue spatial visualization tests. West Lafayette, IN: Purdue Research Foundation, 1977.

Guay, R. Factors affecting spatial test performance: Sex, handedness, birth order, and experience. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978. (a)

Guay, R. The development and concurrent validation of a testing program to select machine maintenance personnel (Report AVN-2). West Lafayette, IN: Applied Management Services, Inc., June 1978. (b)

Guay, R., & McDaniel, E. Correlates of performance on spatial aptitude tests (Report DAHC 19-77-G-0019). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, November 1978.

Guay, R., McDaniel, E., & Angelo, S. Analytic factor confounding spatial ability measurement. Paper presented at the annual meeting of the American Psychological Association, Toronto, August 1978.

Harris, L. Sex differences in spatial ability: Possible environmental, genetic and neurological factors. In M. Kinsbourne (Ed.), Hemispheric Asymmetries of Function. Cambridge, England: Cambridge University Press, 1976.

Kelley, T. Crossroads in the mind of man. Stanford, CA: Stanford University Press, 1928.

Likert, R., & Quasha, W. The revised Minnesota paper form board test manual. New York: The Psychological Corporation, 1970.

Maccoby, E., & Jacklin, C. The psychology of sex differences. Palo Alto, CA: Stanford University Press, 1974.

McGlone, J., & Kertesz, A. Sex differences in cerebral processing of visuo-spatial tasks. Cortex, 1973, 9, 313-320.

Michael, W., Guilford, J., Fruchter, B., & Zimmerman, W. The description of spatial visualization abilities. Educational and Psychological Measurement, 1957, 17, 185-199.

Price, L., & Eliot, J. Convergent and discriminant validities of two sets of measures of spatial orientation and visualization. Educational and Psychological Measurement, 1975, 35, 975-977.

Shepard, R. N., & Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

Smith, I. Spatial ability. San Diego: Knapp, 1964.

Stanard, S., & Bode, K. The SRA test of mechanical concepts. Chicago: Science Research Associates, 1976.

Talsma, G., & Wheatley, G. Assessing spatial ability of young adolescents. Manuscript submitted for publication, 1980.

Thurstone, L., & Thurstone, T. The SRA verbal test. Chicago: Science Research Associates, 1973.

Vandenberg, S. Sources of variance in performance on spatial tests. In J. Eliot & N. Salkind (Eds.), Children's Spatial Development. Springfield, IL: Thomas, 1975.

---

Note. The correct answer for the sample PSVR Test item presented in Figure 2 is B.